Amendments to the Specification:

Please replace the paragraph beginning at page 2, line 25, with the following amended paragraph:

The process of defining a near optimal model in mixed effect non linear regression, non linear regression and logistic regression is commonly called model building. In pharmacokinetic/pharmacodynamic modeling, data from the system of interest (usually a population of patients or normal subjects) is used to estimate the parameters of a mathematical model. Occasionally, attempts are [[make]] made to break the system of interest down into smaller parts, each of which is then used to estimate the parameters of a model. NONMEM is the industry standard software for estimating parameters for a model, given a data set and a "model". The model is a set of equations (algebraic or differential) that are intended to describe the system of interest. Once the set of equations is identified, the parameters of those equations are estimated (by "fitting") by NONMEM or similar software. The "model building" part consists of an often long process of testing various models (sets of equations) for their ability to describe the observed data. The model is then modified, or rejected and a new model is tested. An example is described below.

Please replace the paragraph beginning at page 5, line 3, with the following amended paragraph:

The two compartment system has five parameters, k (mass transfer rate constant out [[or]] of compartment 1), k12 (mass transfer rate constant from compartment 1 to 2), k21 (mass transfer rate constant from compartment 2 to 1), V(1) (volume of compartment 1) and V(2) (volume of compartment 2). Other, more complex system (3, 4 or occasionally more compartments) may be appropriate to describe other drugs. In addition to the selection of the number of compartments, a sub-model may best describe each parameter of the model. For example, it is frequently observed that the volume distribution is well described by a linear function of weight, of the form Volume = $\Theta \bullet$ weight where Θ is a constant. If this is found to be the case, the equation

for a one compartment model becomes:

$$Concentration = \frac{Dose}{\theta \bullet weight} \bullet e^{-kt}$$

Where Θ is a constant describing the relationship between weight and volume.

Please replace the paragraph beginning at page 8, line 1, with the following amended paragraph:

Where Y is the observed value, F is the predicted value Θ is an estimated parameter and ε is a random variable with mean zero. The first three models have a single parameter to be fitted, the variance of ε, the fourth model has two parameters to be [[estimates]] estimated Θ and the variance of ε, and the fifth has two parameters to be fitted, the variance of ε(1) and the variance of ε(2). Additionally, models for autocorrelated residuals can be implemented.

Please replace the paragraph beginning at page 14, line 14, with the following amended paragraph:

It is the object of the present invention to provide improved methods, systems and computer program products for identifying the optimal or near optimal model of the concentrations or pharmacological effects of a drug or drugs. The central concept is to identify a candidate model search space, then search that space. The candidate model search space will be defined as having n dimensions where a dimension is a mutually exclusive set of model features. The dimensions of the search space have discrete values. For example, a [[parametergd]] parameter either is (value = 1) or is not (value = 0) a specific function of a demographic variable (covariate). This dimension has two values, 0 and 1. A value of 1.5 is not possible.

Please replace the paragraph beginning at page 15, line 1, with the following amended paragraph:

a) defining a candidate search space having n dimensions, wherein n is a positive integer and each dimension represents a set of mutually exclusive features from which exactly one of said mutually exclusive features is chosen from each set of mutually exclusive features for each candidate model; and

Please replace the paragraph beginning at page 16, line 1, with the following amended paragraph:

a) defining a candidate search space having n dimensions, wherein n is a positive integer and each dimension represents a set of mutually exclusive features from which exactly one of said mutually exclusive features is chosen from each set of mutually exclusive features for each candidate model; and

Please replace the paragraph beginning at page 18, line 1, with the following amended paragraph:

a) defining a candidate search space having n dimensions, wherein n is a positive integer and each dimension represents a set of mutually exclusive features from which exactly one of said mutually exclusive features is chosen from each set of mutually exclusive features for each candidate model; and

Please replace the paragraph beginning at page 19, line 1, with the following amended paragraph:

a) defining a candidate search space having n dimensions, wherein n is a positive integer and each dimension represents a set of mutually exclusive features from which exactly one of said mutually exclusive features is chosen from each set of mutually exclusive features for each candidate model; and

Please replace the paragraph beginning at page 19, line 10, with the following amended paragraph:

a) defining a candidate search space having n dimensions, wherein n is a positive integer and each dimension represents a set of mutually exclusive features from which exactly one of said mutually exclusive features is chosen from each set of mutually exclusive features for each candidate model; and

Please replace the paragraph beginning on page 20, line 19, with the following amended paragraph:

Figure 3 is as depiction of the 3 dimensional candidate search space described in Figure [[1]] 2, with 27 possible models identified.

Please replace the paragraph beginning at page 25, line 14, with the following amended paragraph:

According to Goldberg "Genetic algorithms are search algorithms based on the mechanics of natural selection and natural genetics".[1] Genetic algorithm is chosen over the other methods for a demonstration of this invention for a number of reasons. Traditional optimization techniques are limited to continuous parameters, that is, the parameters can take any value. The descriptions of a model however, is discrete, either a feature is present in a model or it isn't, and the model may be one compartment or two compartments, but not 1.5 compartments. There are seven methods for optimization of discrete systems. These are:

Please replace the paragraph beginning at page 27, line 12, with the following amended paragraph:

The objective function in NONMEM is a measure of the goodness of fit. This number is equal to –2 times the log likelihood of the observed data, given the model. In addition to the objective function value, which describes the goodness of the fit of the model to the data, parsimonious model are generally preferred, that is, we would like

the simplest model that describes the data well. Therefore, a cost (or penalty) is typically applied for each parameter that is fitted in the model. The user may assign this value, but a commonly used value is 7.84, based on the log likelihood test. Random effects in the model ~~also be addressed~~ are typically addressed in conventional model building. The parameters for one person will vary from those of another person. For example, a parameter might be weight. Typically, weight can be directly measured, but assume for a moment that we are trying to estimate as a parameter of a model. Weight will vary from one person to another, with some population mean and standard deviation. The mean and standard deviation of this distribution is a random effect model. Further, the shape of the distribution is specified in the model. Shapes of distributions include normal (Gaussian), log-normal, beta etc. These again, are discrete features that might be included in the models. The Akaike information criterion suggests that a value of 2 may be appropriate for each random effect entered into the model.

Please replace the paragraph beginning at page 28, line 11, with the following amended paragraph:

In addition there are several other desirable attributes of a model fit. First, that the minimization conclude successfully. That is, the requested number of significant digits is obtained. Second, that the covariance step be executed successfully, so that standard errors of the estimate can be obtained. Finally, the estimation correlations between parameters are all less than 0.95. The user of the algorithm can enter value for the penalty for these. For an adequate model, all these ~~attributed~~ attributes are typically required to be present. Therefore, a large penalty for each of these (~400) will typically be used.

Please replace the paragraph beginning at page 29, line 11, with the following amended paragraph:

A niching penalty can be calculated in a variety of ways, such as fitness sharing and implicit sharing[2]. In this application, we have chosen a novel method. In method, the user defines the number of niches to be defined in the population, and the niche

radius. (niche radius is simply the number of loci that the two individuals differ at). The most fit individual is then selected. All individuals within 1 niche radius of that individual are considered to be in that niche. The next most fit individual, not currently in a niche is then selected, and the next most fit niche is defined as those individuals within one niche radius of that individual. This process is repeated for the number of user defined niches. Those individuals not selected are considered not to be in a niche. Then a user defined "niche penalty" is added to each individual not in a niche. Typically, this niche penalty will be a fraction (perhaps 80%) of the difference between the most fit individual in the niche and the most fit individual that is not in a niche. This niche penalty is then divided between all the members of a given niche.

Please replace the paragraph beginning at page 30, line 25, with the following amended paragraph:

Next, mutations are introduced into the genome strings. The probability of mutation is defined by the user, typically between 0.01 and 0.001. The strings are looped over, a random number between 0 and 1 is generated for each loci on each string. If the value of the random number is less than the probability of mutation, the value at that locus is reversed (1 changed to 0, 0 changed to 1). Finally, a more large-scale change in the genome can be introduced by a frame shift mutation. A frame shift mutation consists of moving the values of all loci in a sub string of the genome one loci left or right. Given that this dramatically changes the resulting model, the probability of it occurring is typically very low (0.01). If an individual is (randomly) selected two loci in that string are randomly selected. The values between those loci are shifted left to right or right ~~or~~ to left, depending on whether the first or second loci is to the left.

Please replace the paragraph beginning at page 38, line 23, with the following amended paragraph:

This is applied to the search for an optimal model as follows: An initial random model is created. An initial high temperature is defined. The "energy" is calculated. The energy in simulated annealing is analogous to the fitness in genetic algorithm, except we want to minimize the energy and maximize the fitness. A random change is

introduced into the model. This may be by change of the value in one or more dimensions. The energy of the new model is calculated. If the energy of the new model is lower (the model is better), the change is retained. If the energy of the new model is higher, the model may be retained. If the energy is higher, the new model is retained with probability:

Please replace the paragraph beginning at page 39, line 4, with the following amended paragraph:

Note that if ΔE is negative, the value of this expression is greater than one, and the change will always be retained. If ΔE is positive (the new model is not as good as the former model), the change may still be retained, depending on the value of ΔE and T. As T decreases, the probability of retaining a change that results in a worse model ~~increases~~ decreases, the model becomes "frozen".

Please replace the paragraph beginning at page 40, line 4, with the following amended paragraph:

A library that implements tabu search and scatter search/path relinking is commercially available (OptQuest callable library from OptTek Systems Inc, Boulder, CO). The implementation of this would be very similar to the implementation of genetic algorithm[[.]], except that a call to the function OCLGetSolution would be made to generate the next individual and the resulting fitness would be added to the data set using the function OCLPutSolution.